J Mol Evol (2000) 50:31–38
DOI: 10.1007/s002399910004

JOURNAL OF MOLECULAR EVOLUTION

© Springer-Verlag New York Inc. 2000

Expression of the Dentin Matrix Protein 1 Gene in Birds

Satoru Toyosawa,* Akie Sato, Colm O'hUigin, Herbert Tichy, Jan Klein

Max-Planck-Institut für Biologie, Abteilung Immungenetik, Correnstrasse 42, 72076 Tübingen, Germany

Received: 5 April 1999 / Accepted: 9 August 1999

Abstract. The emergence of jawed vertebrates was predicated on the appearance of several innovations, including tooth formation. The development of teeth requires the participation of several specialized genes, in particular, those necessary for the formation of hard tissues—dentin, enamel, and cementum. Some vertebrates, most conspicuously birds, secondarily lost the tooth-forming ability. To determine the fate of some of the tooth-forming genes in the birds, we tested a domestic fowl cDNA library for the expression of the dentin matrix protein 1 (*DMP1*) gene. The library was prepared from the poly(A^+) RNA isolated from the jaws of 11- to 13-day-old embryos and the testing was carried out by the polymerase chain reaction with degenerate primers designed on the basis of the available mammalian and reptile sequences. A chicken homologue of the *DMP1* gene identified by this approach was shown to be expressed in the jaws and long bones, the same two tissues as in mammals. The chicken *DMP1* gene has an exon/intron organization similar to that of its mammalian and reptile counterparts. The chicken gene contains three short highly conserved segments, the rest of the gene being poorly alignable or not alignable with its mammalian or reptilian homologues. The distribution of similarities and dissimilarities along the gene is indicative of a mode of evolution in which only short segments are kept constant, while the rest of the gene is relatively free to vary as long as the proportion of certain amino acid residues is retained in the encoded polypeptide. The *DMP1* gene may have been retained in birds because of its involvement in bone formation.

Key words: Tooth formation — Dentin matrix protein 1 — *DMP1* — Hen's tooth

Introduction

Dentin is composed of mineralized extracellular matrix secreted by odontoblasts (Butler and Ritchie 1995). The matrix consists of collagen (about 90%) and noncollagenous components (about 10%), mostly acidic phosphoproteins. The collagen fibrils function as a scaffolding for the deposition of hydroxyapatite crystals, whereas the phosphoproteins are believed to be involved in the initiation of mineralization and in the control of the rate of apatite growth. The best characterized of the phosphoproteins is encoded in the dentin matrix protein 1 (*DMP1*) gene, which has been cloned in human (Hirst et al. 1997b), cattle (Hirst et al. 1997a), mouse (MacDougall et al. 1998), rat (George et al. 1993), wallaby (Toyosawa et al. 1999a), opossum (Toyosawa et al. 1999a), platypus (Toyosawa et al. 1999a), and caiman (Toyosawa et al. 1999b). Transcripts of the *DMP1* gene were originally detected by northern hybridization in rat odontoblasts (George et al. 1993) but were later shown by in situ hybridization to be present in other mineralized tissues, including enamel, cementum, and bone (D'Souza et al. 1997; George et al. 1994), as well as in fetal cattle brain (Hirst et al. 1997a). The human *DMP1* gene codes for a polypeptide chain composed of 16 amino acid residues of a signal sequence and 473 residues of the mature protein (Hirst et al. 1997b). The acidity of the protein is

* *Present address:* Department of Oral Pathology, Osaka University, Faculty of Dentistry, 1-8, Yamadaoka, Suita, Osaka 565-0871, Japan
Correspondence to: Jan Klein; *e-mail:* jan.klein@tuebingen.mpg.de

attributed to the preponderance of Asp and Glu residues, and its phosphoprotein nature to the presence of >100 Ser residues forming sites for phosphorylation by casein kinases I and II (Butler and Ritchie 1995; Hirst et al. 1997b). The protein sequence contains several potential *N*- and *O*-glycosylation sites. The *DMP1* presumably functions by binding to the matrix fibrils and thus initiating nucleation and facilitating mineralization through its high Ca^{2+}-binding capacity.

Since the *DMP1* gene is present in proto-, meta-, and eutherian mammals, as well as in crocodilian reptiles (Hirst et al. 1997a,b; MacDougall et al. 1998; George et al. 1993; Toyosawa et al. 1999a,b), it was presumably also possessed by the toothed ancestors of the birds. Possible retention of at least some tooth-forming capacity by birds is indicated by tissue recombination experiments. From these, it has been claimed that the *in vitro* cocultivation of a neural crest-derived mesenchyme of a mouse (Kollar and Fisher 1980) or lizard (Lemus et al. 1983; Lemus 1995) embryo with the oral epithelium of a fowl or quail embryo leads to the production of amelogenin by bird-derived ameloblasts whose differentiation has been stimulated by inductive signals from the mammalian tissue (Lemus 1995). A reciprocal experiment leading to the purportedly chicken-derived dentin has also been reported (Lemus 1995), but not all studies have supported the view that chicken tooth-forming genes are potentially functional (Karcher-Djuricic and Ruch 1973; Arechaga et al. 1983). To determine whether any of the genes directly involved in tooth formation are expressed in birds, we attempted the isolation of the bird *DMP1* gene.

Materials and Methods

Source and Isolation of DNA

Domestic fowl (*Gallus gallus*) embryos at different days (11 through 13) of development were collected, their upper and lower jaws, as well as other parts, were dissected out, and the tissue blocks were immediately frozen in liquid nitrogen for later use in the preparation of the cDNA library. Adult Reeves's pheasant (*Syrmaticus reevesii*) tissue was used for the isolation of genomic DNA by the phenol–chloroform extraction method (Davis et al. 1986).

Fowl cDNA Library

Poly(A^+) RNA was isolated using the mRNA purification kit and cDNA was synthesized with the help of the TimeSaver cDNA synthesis kit (Pharmacia Biotech). The cDNA was then inserted into *Eco*RI-digested λgt10 vector (Stratagene), in vitro-packaged with the help of the Gigapack cloning kit (Stratagene), and used to transform competent *E. coli* NM514 bacteria. The cDNA libraries were screened by polymerase chain reaction (PCR) amplification using degenerate primers based on an alignment of human, cattle, rat, mouse, wallaby, opossum, platypus, and caiman *DMP1* sequences. The two primers were as follows: CD10, sense, 5′-GAGGAGGAYGAYAGYGGNGAYGAYAC-3′; and CD8, antisense, 5′-ATAACCATCYTGRCARTCRTTRT-CRTC-3′. The amplification product of these two primers spans a region of approximately 1 kb in exon 6 of the *DMP1* gene.

PCR Amplifications

Genomic DNA (100 ng/μl) or lysate of the cDNA libraries (1 μl) were amplified by PCR in 50 μl of PCR buffer (1.5 m*M* $MgCl_2$/200 μ*M* dNTP/10 m*M* Tris, pH 8.5) in the presence of the sense and antisense primers and 2.5 U of *Taq* polymerase (Pharmacia Biotech). Amplifications were performed in the RoboCycler Gradient 96 (Stratagene) in 35 cycles, each cycle consisting of 1 min of denaturation at 94°C, 1 min of annealing at the annealing temperature, and 3 min of extension at 72°C. The final extension was for 10 min at 72°C. The annealing temperatures varied from 53 to 60°C depending on the primer combination used. For the initial isolation with the CD10-CD8 primer combination, the temperature was 57°C. For reverse transcription PCR, 1 μg of total RNA prepared from 11- to 13-day-old jaws of fowl embryos was used for the first-strand cDNA synthesis with random hexamer primers, aided by the 5′ RACE kit (GIBCO/BRL). The product was amplified by PCR using the fowl *DMP1*-specific primer DEN95 (antisense, 5′-GCTGCTGCTGCTACTGTCACCATGTC-CTCT-3′) in combination with the generic anchor primer supplied with the kit and annealing to the poly(C) tail. The DEN95 primer anneals to exon 6 of the *DMP1* gene, 342 bp from the cDNA's start codon. The PCR product was subcloned into *Sma*I-digested pUC18 plasmid vector.

DNA Sequencing

Double-stranded DNA was purified using the Qiagen plasmid kit. Two to five micrograms of DNA was used in the dideoxy chain-termination reactions performed with the help of the AutoRead Sequencing Kit (Pharmacia Biotech), which contained 5′-fluorescent-labeled sequencing primers and T7 DNA polymerase. The reactions were then processed by the Automated Laser Fluorescent (ALF) Sequencer (Pharmacia Biotech). The following two internal primers deduced from the *DMP1* sequence were used: DEN96, sense, 5′-CCAGAGACTGA-GGACACCAACTCT-3′; and DEN97, antisense, 5′-TTGGTGC-TCTCTGTGGATGTGCT-3′. To determine the exon–intron organization of the fowl *DMP1* gene, the following primers were used to amplify the individual boundary regions (Fig. 1): DEN98, sense (E2), 5′-ATGAAGACTGCACTGCTGATGTTG-3′; DEN99, antisense (E3), 5′-AGGCTCTCGGGACACGGAGAGATG-3′; DEN100, sense (E3), 5′-CATCTCTCCGTGTCCCGAGAGCCT-3′; DEN101, antisense (E6), 5′-GGGACCCAGCGCATTTGCCACATC-3′. Similarly, the primers for the exon–intron boundaries of the pheasant gene were as follows: DEN104, sense [E2, 5′ untranslated region (UTR)], 5′-ACGCTTCTACACCTCTGCTGCAGG-3′; and DEN113, antisense (E6), 5′-ACCATGTCCTCTGGTGTCCCCAGT-3′. The DEN104 primer, in combination with DEN105, antisense (E6, 3′UTR), 5′-TGACCCAACTCCACCACAAACAAA-3′, was used to confirm the identity of sequences in the overlapping parts of the cDNA clones. To obtain the pheasant *DMP1* sequence from the genomic DNA, the following three primer pairs were used: DEN104/DEN101, DEN100/DEN113, and CD10/CD8.

Fig. 1. Nucleotide and translated amino acid sequence of fowl *DMP1* cDNA clones. Exon/intron borders are indicated by the ↓ symbols. Amino acid residues are given in the IUPAC–IUB single-letter code. Primer positions and orientations are indicated by an *arrowed line* above their location in the sequence.

DEN104
GGACGAGCACCACACGCTTCTACACCTCTGCTGCAGG

DEN99/DEN100

DEN98

1 11 ↓ 21
ATG AAG ACT GCA CTG CTG ATG TTG CTG CTG TGG GCT GCA GCC CGC GCC CAC CCT GTG CCT GGC CAT CTC TCC GTG
M K T A L L M L L L W A A A R A H P V P G H L S V

26 ↓ 36 DEN101 46
TCC CGA GAG CCT GAG GAT GCA GCA CCA GGG GAT GTG GCA AAT GCG CTG GGT CCC CTC CTG GGT GAT GAA GAC CAT
S R E P E D A A P G D V A N A L G P L L G D E D H

DEN112
51 61 71
GGG CGC CTC CAT GCC AGC ATG GGG CCA TGG GAC ACT GGT GAT GCG GAC GTA GCC TTT GGG AAC CAC GTG GAG CGA
G R L H A S M G P W D T G D A D V A F G N H V E R

CD10
76 86 96
GAC CTT GCT GAG CAC AAC GGG CTG CCT GCC TGG CCT GGG GAC GAG GAG GAT GAC AGC GGG GAT GAC ACC TTT GAT
D L A E H N G L P A W P G D E E D D S G D D T F D

DEN95

DEN113
101 111 121
GAG GAT AAG GAA GAG GGT GAA GGG CCT GCT TAC GGC ACT GGG GAC ACC AGA GGA CAT GGT GAC AGT AGC AGC AGC
E D K E E G E G P A Y G T G D T R G H G D S S S S

DEN115
126 136 146
AGC AGC AGT GAG AGT GCG ATG GCC CCA CGC CGC CAC ATC CGA TAC CGT GGC GGC TCC AGG TGG GGA GGA GGC AGC
S S S E S A M A P R R H I R Y R G G S R W G G G S

151 161 171
AGC AGC AGC CAG GAG GAT GAT GAG GAG AAC ATT GGA CTG GGC AGT GAA GGC ATG CAG GGT GAC GAC CCC TCT GTC
S S S Q E D D E E N I G L G S E G M Q G D D P S V

176 186 196
TTT GAC AGC CTG GGT GGC AAG CAC CGC AGC CGG GGG ACG TCT GGC AGC CCC TGG GAG GAT GAA GAC AGC CAC TCC
F D S L G G K H R S R G T S G S P W E D E D S H S

DEN96
201 211 221
CCA GAG ACT GAG GAC ACC AAC TCT GTG GAA GAC AGC AAG GAG AAG AGT CAC TCA CAG GAG AAC AGC GAG GCC AGC
P E T E D T N S V E D S K E K S H S Q E N S E A S

226 236 246
AGG TCC AGG GAG GAT GGC GAC AGC CCA TCC CAG GAG GAT GGG GAC AGT CCG TCC AAA GAG GAT GGG GAC AGC CCG
R S R E D G D S P S Q E D G D S P S K E D G D S P

251 261 271
TCC CAG GAG GAT GCG GAC AGC CCG TCC CAG GAG GAT GCG GAC AGC CCA TCG CAG GAG GCA TCA GAT GAG GAG TCT
S Q E D A D S P S Q E D A D S P S Q E A S D E E S

276 286 296
GCA GAG GAC GGC TCA GAG GAG GCG GTG AGC GCA TCC CAG GGA CGC AGC AGC AGG GAG GAG CGG GCG TCC GCA GAG
A E D G S E E A V S A S Q G R S S R E E R A S A E

301 311 321
GAC AGA AGC GCA GTG TCT GAT CTT GAC ACC GAG GAG GAG CAG AGC AGG TCC AAA GAG GAC AGC TTG GAG ACA GAG
D R S A V S D L D T E E E Q S R S K E D S L E T E

326 336 346
GAA GAT GTC TCC AAG CCT GAC GAT GAT GCT CCA AGT GCA TCA ACT GAG AGC CAG AGC TCA TCC CCA GAG GGC AGC
E D V S K P D D D A P S A S T E S Q S S S P E G S

DEN97
351 361 371
CAG GAG GAC AGT GAG GGT GAG GAG GAC GAG GCA GTG AGC GAA GAG AGC ACA TCC ACA GAG AGC ACC AAC AGC GCA
Q E D S E G E E D E A V S E E S T S T E S T N S A

376 386 396
TCC CTG GAG GAG GAT GAT GAT GAT GAT GTG CAC TCC CAG GAG ACT ACC AGC CGT GGG GAT GCC AGC TCA CTG CGC
S L E E D D D D D V H S Q E T T S R G D A S S L R

CD8
401 411 421
AGC CTG AAG AAC CGC AGG CGG CGG CCG GGT GCC TAC CAC AGG AAG CAG GCT GCC GAG CTC GAC GAC GAC GAC TGC
S L K N R R R R P G A Y H R K Q A A E L D D D D C

426
CAG GAT GGG TAC TGA
Q D G Y Ter

DEN105
TGGCACTGCTGCGTCATTTGTTTGTGGTGGAGTTGGGTCAGCCTCTGCTTCTGCGGACAGACAGGGTGAGGGCTGGCCATGTGGGCAGTGAGTGGTGGGG
CTGGCTGTGGGGAATCCCAAGTCCTGGCCTGCAGGTGAGGGCATCCCAGCGCTCTCTCCTTGCCCTTGCTGCTGGGAACCGCACACGGTATCATGCTGGC
GCACATCGCACCACGCTGAGGGGGGGGGGGGGTGTATGTAATATTTTGTAAGAAATGTACAGAAAGCCTTCTATTGTAACCTCTCTGTGTCCACGGCACT
TATCTCTCCATGCTGAAATGTATCCTGTGTATGTCCTTGATTCACATTAAACA

Southern and Northern Blot Analysis

Southern blotting was carried out as described earlier (Toyosawa et al. 1999b). For northern blotting, total RNA and poly(A^+) RNA were prepared using the Qiagen RNeasy and the Oligotex mRNA kits, respectively. The mRNA was electrophoretically separated on formaldehyde gel and blotting was performed (Davis et al. 1986). The blots were hybridized according to the protocol provided by the supplier of the Alk-Phos Direct kit (Amersham) used for labeling the probe.

Sequence Analysis

Alignments of protein sequences were made with the aid of the ClustalW program (Thompson et al. 1994) and corrections were made by hand. For detailed analysis of conservation patterns, the fowl, pheasant, caiman, mouse, and human sequences were separately aligned, following omission of 115 amino acid extension in exon 5, unique to caiman (Toyosawa et al. 1999b). A sliding window of 30 codons was used to estimate the percentage of nonsynonymous substitution (K_a) for pairwise comparisons by the method of Li (1993).

Results and Discussion

To determine the fate of the *DMP1* gene in modern birds, we prepared cDNA libraries using mRNA isolated from the jaws of domestic fowl (chicken) embryos at 11–13 days of development and screened them by PCR using degenerate primers based on the sequences of mammalian and reptilian *DMP1* genes. The presence of positive clones in the cDNA libraries was indicated by strong PCR signals. The positive clones were isolated and sequenced. The complete coding sequence of the fowl cDNA was obtained from three overlapping clones. One of the clones (dmp1-1), which contained the entire translated region, was PCR-amplified from the cDNA library using the primer pair CD10/CD8 (see Materials and Methods). The second cDNA clone (dmp1-2), which contained the 3′ UTR, was amplified by anchored PCR using a vector-specific primer and the internal primer DEN91. The third clone, encompassing a portion of the 5′ UTR, was obtained by reverse transcription PCR. The primer pair DEN104/DEN105 was then used to confirm the identity of the entire coding sequence (Fig. 1).

The 1680-bp-long sequence was found to contain an open reading frame for 429 amino acid residues, beginning with the initiation codon at nucleotide 38. The first 21, mostly hydrophobic, amino acid residues of the deduced protein sequence presumably constitute the signal peptide, followed by 408 residues of the mature protein. The stop codon at nucleotide site 1325 is followed by 353 bp of the 3′ UTR. Exon–intron boundaries of the fowl *DMP1* gene were determined by sequencing the corresponding regions of genomic DNA, using the known genomic organization of the reptile and mammalian genes as a guide. The fowl *DMP1* gene consists of four exons interrupted by three introns, two fewer than in the mammalian *DMP1* genes (Fig. 2). An exon corresponding to the mammalian exon 1 and containing the 5′ UTR presumably exists in the fowl gene but it could not be identified in the cDNA clones, perhaps because of the way the cDNA libraries were prepared. The fowl *DMP1* exon 2 corresponds exactly to the mammalian and reptilian exon 2. The 3′ border of fowl exon 3 is shifted by four codons upstream compared to the mammalian exon 3. Mammalian/reptilian exons 4 and 5 appear to be missing in the fowl *DMP1* gene. The 5′ border of fowl exon 6 is shifted downstream compared to both the mammalian and the reptilian *DMP1* genes (the 5′ border of the mammalian gene being shifted downstream relative to the reptilian gene). Both variation in exon size and differential splicing of the upstream exons of *DMP1* have been described previously (MacDougall et al. 1998; Toyosawa et al. 1999, a,b). Exon 5 is differentially spliced in the mouse *DMP1* (MacDougall et al. 1998) and can vary eightfold in size, from 45 bp in the mouse to 357 bp in the caiman (Toyosawa et al. 1999, a,b). The functional significance of the absence of exons 4 and 5 in the fowl remains unclear. The exons are moderately conserved in eutherian mammals (Fig. 3), which suggests that they are constrained in their evolution. However, the observed variation of exon 5 and the absence of exons 4 and 5 in some birds imply that these exons play a role which is not critical to the function of the DMP1 protein and that their absence may not impair the function. The conservation of fowl and pheasant *DMP1* gene organization suggests that exons 4 and 5 were absent before the splitting of the two lineages, some 16–20 million years ago (Helm-Bychowski and Wilson 1986).

The comparison of the fowl deduced protein sequence with the known *DMP1* sequences reveals the existence of three short, highly conserved regions, the rest of the protein being poorly alignable with either the mammalian or the reptilian polypeptides (Fig. 3). The conserved regions span residues 218–226, 328–335, and 655–663. The remainder of the sequence, although highly divergent among mammals, reptiles, and birds, nevertheless retains a similar character in terms of amino acid composition. The putative fowl DMP1 protein, like the known DMP1 molecules of other species (Hirst et al. 1997a,b; MacDougall et al. 1998; George et al. 1993; Toyosawa et al. 1999a,b), has a relatively high content of acidic amino acids Asp (13.4%) and Glu (13.5%), the two residues responsible for the acidic character of the polypeptides, as well as a high content of Ser (17.5%), the residue essential for the phosphorylation of the molecule.

This distribution of similarities and dissimilarities along the sequences is indicative of a mode of evolution in which only short segments are kept constant, while the rest of the polypeptide is relatively free to vary, as long as the proportion of a few specific amino acid residues is retained. This unusual mode of evolution becomes apparent in a sliding window comparison of the percentage

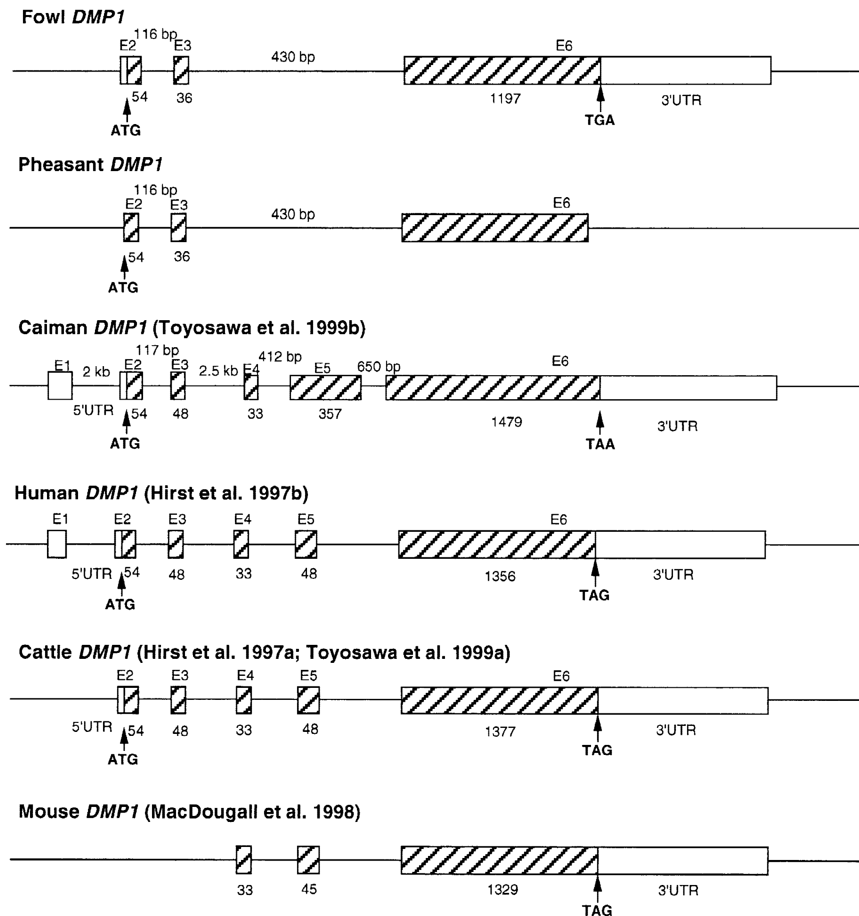

Fig. 2. Exon (E)–intron organization of *DMP1* genes in birds, a reptile, and mammals. *Hatched* and *open rectangles* indicate translated and untranslated (UTR) exon regions, respectively. Coding lengths of exons in base pairs are indicated beneath the boxes; lengths of introns in base pairs or kilobase pairs are also indicated, where known. ATG and TGA or TAA are the initiation and the stop codons, respectively. Sources of information are referenced after the species names.

of nonsynonymous substitutions per codon positions in the genes of the different pairs of taxa (Fig. 4). The comparison reveals that nonsynonymous sites diverging rapidly between closely related taxa (representatives of the same or different mammalian orders) correspond to the nonconserved sites in comparisons between distantly related taxa (different vertebrate classes). Similarly, the slowly evolving nonsynonymous sites in intra- or interordinal comparisons within the same class correspond to the highly conserved sites in comparisons of bird, reptile, and mammalian sequences. The expectation is that a similar mode of evolution should also be apparent in comparisons of *DMP1* genes from different species of birds. This expectation is borne out by the *DMP1* sequence of the Reeves's pheasant (Fig. 3), which last shared a common ancestor with the domestic fowl approximately 16–20 million years ago (Helm-Bychowski and Wilson 1986). The pheasant and the fowl sequences are alignable along their entire length (Fig. 3) but have clearly diverged more in the segments that cannot be aligned in comparisons of sequences from different vertebrate classes than in the conserved segments. This result also indicates that the bird *DMP1* genes evolve under similar evolutionary pressures as the mammalian and reptilian sequences and hence that they are functional.

The possibility existed, however, that the bird genes did not come from the same, but rather from paralogous, loci within one family. If this were the case, the interpretation of the above results would not be as straightforward as presented. To test for the presence of paralogous or related genes, we performed Southern blot analysis of fowl genomic DNA digested with *Hin*dIII, *Taq*I, *Pst*I, and *Alu*I restriction enzymes and hybridized with two kinds of fowl *DMP1* probes. One probe covered most of exon 6 including the three conserved regions. A second, smaller probe (~200 bp) was obtained by PCR amplification with the primers DEN112 (sense) 5′-AACCACGTGGAGCGAGACCTTGCT-3′ and DEN115 (antisense) 5′-GGAGCCGCCACGGTATCGGATGT-3′. The amplification product encompassed codons 71 to 144 of exon 6 and included only one of the three conserved regions. The hybridization pattern obtained by both probes was identical: only one hybridizing band could be detected (Fig. 5A). A similar result was also obtained with reptile and mammalian genomic DNA and the corresponding probes (Toyosawa et al. 1999a,b). If there were a multigene family of *DMP1*-like sequences sharing the conserved regions, multiple hybridizing bands would be expected to appear on at least some of these blots. The result in Fig. 5A is consistent with the existence of a single *DMP1* gene in the chicken haploid genome.

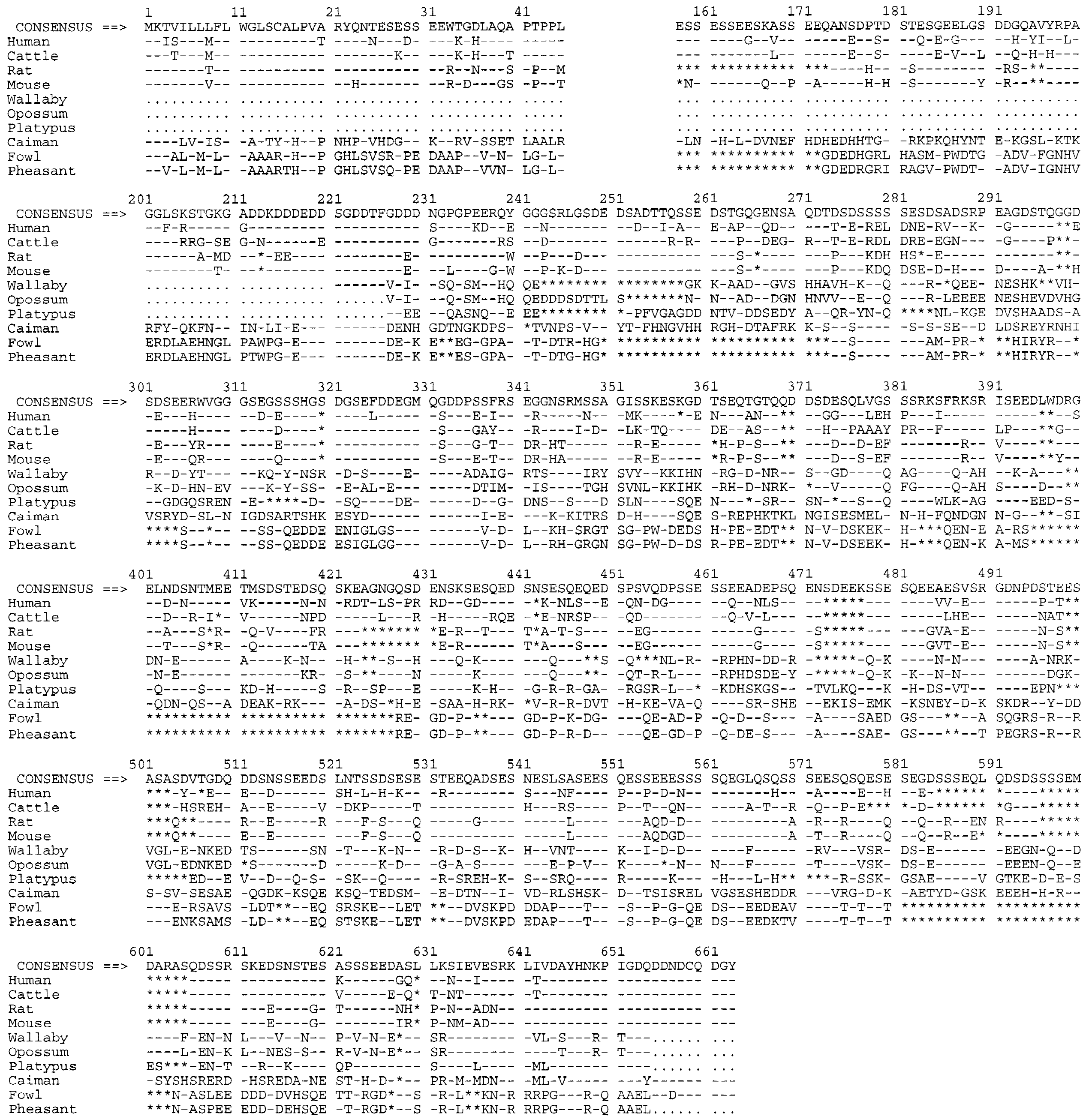

Fig. 3. Amino acid sequence alignment of mammalian, avian, and reptilian DMP1s. The stretch between position 46 and position 158, which is present only in the caiman protein, is omitted. *Dashes* indicate identity with the simple majority consensus sequence at the top; *asterisks,* gaps introduced to optimize the alignment; and *dots,* lack of information. Amino acid residues are given in the IUPAC–IUB single-letter code. Exon boundaries are not considered in the alignment.

Further support for the conclusion that the fowl sequence comes from the *DMP1* gene is provided by the results of northern blot hybridization. Messenger RNA was extracted from the jaws, brain, long (leg) bone, heart, liver, and stomach of 11- and 13-day-old fowl embryos and subjected to northern blot analysis, using ~1.3-kb-long fowl cDNA probe encompassing almost the entire *DMP1* coding region. A single band ~1.7 kb in length was observed only in the lanes containing mRNA derived from the jaw and from the long bone (Fig. 5B). Although we cannot exclude the possibility that the *DMP1* gene is also expressed weakly in other tissues (no additional positive tissues were detected by the PCR amplification of the *DMP1* transcript, however), its main loci of expression in the fowl embryos are the jaws and long bones, the same tissues as in mammals. Finer resolution and identification of the cells expressing the gene must await the results of in situ hybridization of tissue sections. The results of the expression studies are significant in light of what is known about the process of tooth development in mammals. The first step in this process is the migration of neural crest cells into the

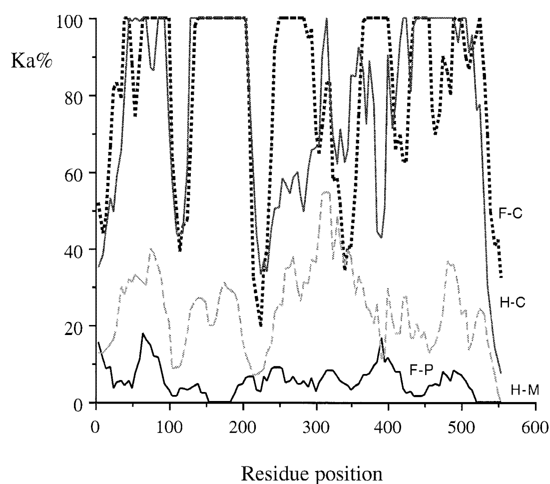

Fig. 4. Plot of percentage nonsynonymous substitutions (K_a%) against codon position for selected pairwise comparison of fowl (F), pheasant (P), caiman (C), mouse (M), and human (H) *DMP1* sequences. A sliding window of 30 codons (including alignment gaps) was used to generate the plot, with measurements taken at 5-codon intervals.

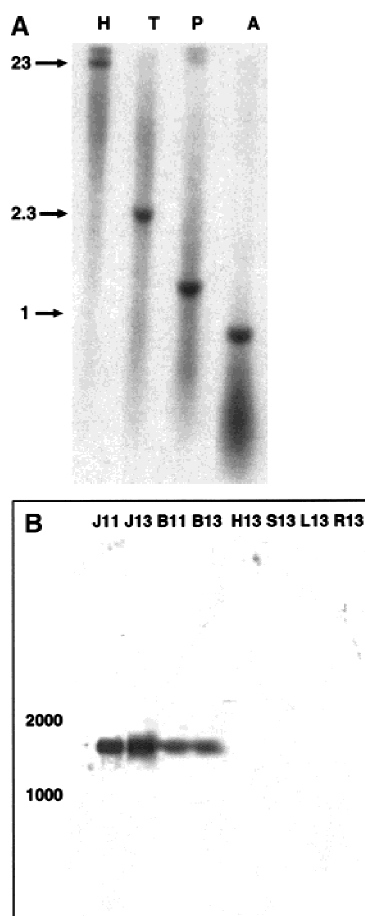

Fig. 5. **A** Southern blot of fowl genomic DNA digested with *Hin*dIII (H), *Taq*I (T), *Pst*I (P), and *Alu*I (A) restriction endonucleases and hybridized with an exon 6 fowl probe covering codons 71–144. The probe was obtained by PCR amplification with the primer pair DEN112/DEN115. A similar pattern was obtained using a probe amplified with the primer pair CD8/CD10. Approximate sizes of fragments are indicated in kilobase pairs. **B** Northern blot of poly(A^+) mRNA isolated from jaws (J) and long (leg) bone (B) of 11- and 13-day-old fowl embryos, as well as heart (H), stomach (S), liver (L), and brain (R) of 13-day-old fowl embryos. The probe was near full-length (1.3 kilobase pairs) fowl cDNA, obtained using the DEN104/CD8 primer pair. Marker sizes (in nucleotides) are indicated on the *left.* Approximately 2.5–3 μg of mRNA was loaded on each lane. The positive control (hybridization with a probe specific for the fowl glyceraldehyde-3-phosphate dehydrogenase transcript) is not shown.

tooth-forming region of the jaws, the so-called maxillary and mandibular arches (Smith and Hall 1990). The mesenchymal cells then induce the overlying dental epithelium to form the dental lamina. The subsequent invagination of the dental lamina into the mesenchyme forms the dental organ, around which the mesenchyme condenses into the dental papilla. We have no information whether any part of this process also takes place in the fowl embryo. Since neural crest cells also give rise to bone structures in the jaws, the observed *DMP1* expression in this organ might not be related to tooth formation at all.

From the experiments described herein, we conclude that the gene we have identified is indeed the bird homologue of the mammalian/reptilian *DMP1* gene. We base this conclusion on the following observations. First, the bird gene contains the same three short highly conserved segments characteristic of the mammalian/reptilian *DMP1* genes. Second, the bird sequence is derived from a single-copy gene which does not appear to be a member of a multigene family. Third, in the birds the gene evolves in the same manner as its presumed homologues in mammals and reptiles—slowly in the conserved regions and rapidly in the divergent regions. Fourth, the fowl deduced protein sequence possesses the same features (high content of Asp, Glu, and Ser residues) as its mammalian/reptilian counterparts. Fifth, the chicken gene has an exon/intron organization similar to that of the *DMP1* genes. Finally, the expression pattern of the chicken gene resembles that of the *DMP1* genes. Hence toothless birds have retained a functional gene that in toothed vertebrates is critically involved in tooth formation.

There may be two reasons why they have done so. The *DMP1* gene of all vertebrates may have other functions in addition to its involvement in tooth formation. Indeed some data on its expression in tissues other than developing teeth seem to suggest a role consistent with inhibition of calcification (Lemus et al. 1983). In this scenario, the gene remains under selective pressure because of the other roles it plays. An alternative scenario postulates that the gene has assumed a new function in birds. In both cases, it is possible that the loss of tooth function would produce a change in the evolutionary rate of the gene. We interpret the similarity of the conservation patterns in birds compared to other vertebrates to indicate that the gene's evolution remains relatively unperturbed by the tooth loss.

In conclusion, we have shown that gross changes in

morphological features may have no major effect on genes directly involved in the ontogeny of these features. The loss of teeth in birds has not caused the concomitant loss of the *DMP1* gene. Other genes, such as amelogenin, may have become redundant and nonfunctional. The maintenance of gene function points to the possibility of future evolutionary innovation involving sidelined genes.

Acknowledgements. We thank Ms. Jane Kraushaar for editorial assistance and Dr. Seiko Shintani for assistance with the graphics.

References

Arechaga J, Karcher-Djuricic V, Ruch JV (1983) Neutral morphogenetic activity of epithelium in heterologous tissue recombinations. Differentiation 25:142–147

Butler WT, Ritchie H (1995) The nature and functional significance of dentin extracellular matrix proteins. Int J Dev Biol 39:169–179

Davis LG, Dibner MD, Battey JF (1986) Basic methods in molecular biology. Elsevier, New York

D'Souza RN, Cavender A, Sunavala G, et al. (1997) Gene expression patterns of murine dentin matrix protein 1 (Dmp1) and dentin sialophosphoprotein (DSPP) suggest distinct developmental functions in vivo. J Bone Miner Res 12:2040–2049

George A, Sabsay B, Simonian PAL, Veis A (1993) Characterization of a novel dentin matrix acidic phosphoprotein. Implications for induction of biomineralization. J Biol Chem 268:12624–12630

George A, Gui J, Jenkins NA, Gilbert DJ, Copeland NG, Veis A (1994) In situ localization and chromosomal mapping of the AG1 (Dmp1) gene. J Histochem Cytochem 42:1527–1531

Helm-Bychowski KM, Wilson AC (1986) Rates of nuclear DNA evolution in pheasant-like birds: Evidence from restriction maps. Proc Natl Acad Sci USA 83:688–692

Hirst KL, Ibaraki-O'Connor K, Young MF, Dixon MJ (1997a) Cloning and expression analysis of the bovine dentin matrix acid phosphoprotein gene. J Dent Res 76:754–760

Hirst KL, Simmons D, Feng J, Aplin H, Dixon MJ, MacDougall M (1997b) Elucidation of the sequence and the genomic organization of the human dentin matrix acidic phosphoprotein 1 (DMP1) gene: exclusion of the locus from a causative role in the pathogenesis of dentinogenesis imperfecta type II. Genomics 42:38–45

Karcher-Djuricic V, Ruch JV (1973) Culture d'associations interspécifiques de constituants mandibulaires d'embryons de souris et de poulet. CR Soc Biol 167:915–918

Kollar K, Fisher C (1980) Tooth induction in chick epithelium: Expression of quiescent genes for enamel synthesis. Science 20:993–995

Lemus D (1995) Contributions of heterospecific tissue recombinations to odontogenesis. Int J Dev Biol 39:291–297

Lemus D, Fuenzalida M, Illanes J, Paz de la Vega Y (1983) Ultrastructural aspects of dental tissues and their behavior in xenoplastic association (lizard-quail). J Morphol 176:341–350

Li W (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 36:96–99

MacDougall M, Gu TT, Luan X, Simmons D, Chen J (1998) Identification of a novel isoform of mouse dentin matrix protein 1: Spatial expression in mineralized tissues. J Bone Miner Res 13:422–431

Smith MM, Hall BK (1990) Development and evolutionary aspects of vertebrate skeletogenic and odontogenic tissues. Biol Rev 65:277–373

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Toyosawa S, O'hUigin C, Klein J (1999a) Dentin matrix protein 1 gene of prototherian and metatherian mammals. J Mol Evol 48:160–167

Toyosawa S, O'hUigin C, Tichy H, Klein J (1999b) Characterization of dentin matrix protein 1 gene in crocodilia: Gene 234:307–314